

Wolfgang G. Stock,
Mechtild Stock

Handbook of Information Science

Contents

A. Introduction to Information Science — 1

- A.1 What is Information Science? — 3
- A.2 Knowledge and Information — 20
- A.3 Information and Understanding — 50
- A.4 Documents — 62
- A.5 Information Literacy — 78

Information Retrieval

B. Propaedeutics of Information Retrieval — 91

- B.1 History of Information Retrieval — 93
- B.2 Basic Ideas of Information Retrieval — 105
- B.3 Relevance and Pertinence — 118
- B.4 Crawlers — 129
- B.5 Typology of Information Retrieval Systems — 141
- B.6 Architecture of Retrieval Systems — 157

C. Natural Language Processing — 167

- C.1 n -Grams — 169
- C.2 Words — 179
- C.3 Phrases – Named Entities – Compounds – Semantic Environments — 198
- C.4 Anaphora — 219
- C.5 Fault-Tolerant Retrieval — 227

D. Boolean Retrieval Systems — 239

- D.1 Boolean Retrieval — 241
- D.2 Search Strategies — 253
- D.3 Weighted Boolean Retrieval — 266

E. Classical Retrieval Models — 275

- E.1 Text Statistics — 277
- E.2 Vector Space Model — 289
- E.3 Probabilistic Model — 301
- E.4 Retrieval of Non-Textual Documents — 312

F. Web Information Retrieval — 327

- F.1 Link Topology — 329
- F.2 Ranking Factors — 345
- F.3 Personalized Retrieval — 361
- F.4 Topic Detection and Tracking — 366

G. Special Problems of Information Retrieval — 375

- G.1 Social Networks and “Small Worlds” — 377
- G.2 Visual Retrieval Tools — 389
- G.3 Cross-Language Information Retrieval — 399
- G.4 (Semi-)Automatic Query Expansion — 408
- G.5 Recommender Systems — 416
- G.6 Passage Retrieval and Question Answering — 423
- G.7 Emotional Retrieval and Sentiment Analysis — 430

H. Empirical Investigations on Information Retrieval — 443

- H.1 Informetric Analyses — 445
- H.2 Analytical Tools and Methods — 453
- H.3 User and Usage Research — 465
- H.4 Evaluation of Retrieval Systems — 481

Knowledge Representation

I. Propaedeutics of Knowledge Representation — 501

- I.1 History of Knowledge Representation — 503
- I.2 Basic Ideas of Knowledge Representation — 519
- I.3 Concepts — 531
- I.4 Semantic Relations — 547

J. Metadata — 565

- J.1 Bibliographic Metadata — 567
- J.2 Metadata about Objects — 586
- J.3 Non-Topical Information Filters — 598

K. Folksonomies — 609

- K.1 Social Tagging — 611
- K.2 Tag Gardening — 621
- K.3 Folksonomies and Relevance Ranking — 629

L. Knowledge Organization Systems — 633

- L.1 Nomenclature — 635
- L.2 Classification — 647
- L.3 Thesaurus — 675
- L.4 Ontology — 697
- L.5 Faceted Knowledge Organization Systems — 707
- L.6 Crosswalks between Knowledge Organization Systems — 719

M. Text-Oriented Knowledge Organization Methods — 733

- M.1 Text-Word Method — 735
- M.2 Citation Indexing — 744

N. Indexing — 757

N.1 Intellectual Indexing — 759

N.2 Automatic Indexing — 772

O. Summarization — 781

O.1 Abstracts — 783

O.2 Extracts — 796

P. Empirical Investigations on Knowledge Representation — 807

P.1 Evaluation of Knowledge Organization Systems — 809

P.2 Evaluation of Indexing and Summarization — 817

Q. Glossary and Indexes — 827

Q.1 Glossary — 829

Q.2 List of Abbreviations — 851

Q.3 List of Tables — 854

Q.4 List of Figures — 855

Q.5 Index of Names — 860

Q.6 Subject Index — 879

A.1 What is Information Science?

Defining Information Science

What is information science? How can we define this budding scientific discipline? In order to describe and delineate information science, we will examine several approaches and pursue the following lines of questioning:

- What are the fundamental determinants and tasks of information science?
- What roles are played by knowledge, information and documents?
- How has our science developed? Do the branches of information science look back on different histories?
- To which other scientific disciplines is information science closely linked?

There is no generally accepted definition of information science (Bawden & Robinson, 2012; Belkin, 1978; Yan, 2011). This is due, first of all, to the fact that this science, at around fifty years of age, is still relatively young when compared to the established disciplines (such as mathematics or physics). Secondly, information science is strongly interrelated with other disciplines, e.g. information technology (IT) and economics, each of which place great emphasis on their own definitions. Hence, it is not surprising that information science today is used for the divergent purposes of foundational research on the one hand, and applied science on the other.

Let us begin with a working definition of “information science”!

Information Science studies the representation, storage and supply as well as the search for and retrieval of relevant (predominantly digital) documents and knowledge (including the environment of information).

Information science is a fundamental part of the development of the knowledge society (Webster, 1995) or the network society (Castells, 1996). In everyday life (Catts & Lau, 2008) and in professional life (Bruce, 1999), information science finds its expression in the skills of information literacy. “We already live in an era with information being the sign of our time; undoubtedly, a science about this is capable of attracting many people” (Yan, 2011, 510).

Information science is strongly related to both computer science (since digitally available knowledge is principally dependent upon information-processing machines) and to the economic sciences (knowledge is an economic good, which is traded or—sometimes—distributed for free on markets), as well as to other neighboring disciplines (such as library science, linguistics, pedagogy and science of science).

We would like to inspect the determinants of our definition of *information science* a little more closely.

- *Representation*: Knowledge contained in documents as well as the documents themselves (let us say, scientific articles, books, patents or company publications, but also websites or microblog posts) are both condensed via short content-

descriptive texts and labeled with important words or concepts with the purpose of information filtering.

- *Storage and supply*: Documents are to be processed in such a way that they are ideally structured, easily retrievable and readable and stored in digital locations, where they can be managed.
- *Search*: Information science observes users as they satisfy their information needs, it observes their query formulations in search tools and it observes the way they use the retrieved information.
- *Retrieval*: The focal points of information science are systems for researching knowledge; prominent examples include Internet search engines, but also library catalogs.
- *Relevance*: The objective is not to find “any old” information, but only the kind of knowledge that helps the user to satisfy his information needs.
- *Predominantly digital*: Since the advent of the Internet and of the commercial information industry, large areas of human knowledge are digitally available. Even though digital information represents a core theme of information science, there is also room left for non-digital information collections (e.g. in archives and libraries).
- *Documents*: Documents are texts and non-textual objects (e.g., images, music and videos, but also scientific facts, economic objects, objects in museums and galleries, real-time facts and people). And documents are both physical (e.g., printed books) and digital (e.g., blog posts).
- *Knowledge*: In information science, knowledge is regarded as something static, which is fixed in a document and stored on a memory. This storage is either digital (such as the World Wide Web), material (as on a library shelf) or psychical (like the brain of a company employee). Information, on the other hand, always contains a dynamic element; one informs (active) or is informed (passive). The production and the use of knowledge are deeply embedded in social and cultural processes; so information science has a strong cultural context (Buckland, 2012). The possible applications of information science research results are manifold and can be encountered in a lot of different places on the information markets (Linde & Stock, 2011). By way of example, we will emphasize six types of application:
 - Search engines on the Internet (a prominent example: Google),
 - Web 2.0 services (such as YouTube for videos, Flickr for images, Last.fm for music or Delicious for bookmarks of websites),
 - Digital library catalogs (e.g. the global project WorldCat, or catalogs relating to precisely one library, such as the catalog of the Library of Congress),
 - Digital libraries (catalogs *and* their associated digital objects, such as the Perseus Digital Library for antique documents, or the ACM Digital Library for computer science literature),

- Digital information services (particularly specialist databases with a wide spectrum reaching from business and press information, via legal information, all the way to information from science, technology and medicine),
- Information services in corporate knowledge management (the counterparts of the above-mentioned five cases of Internet application, transposed to company Intranets).

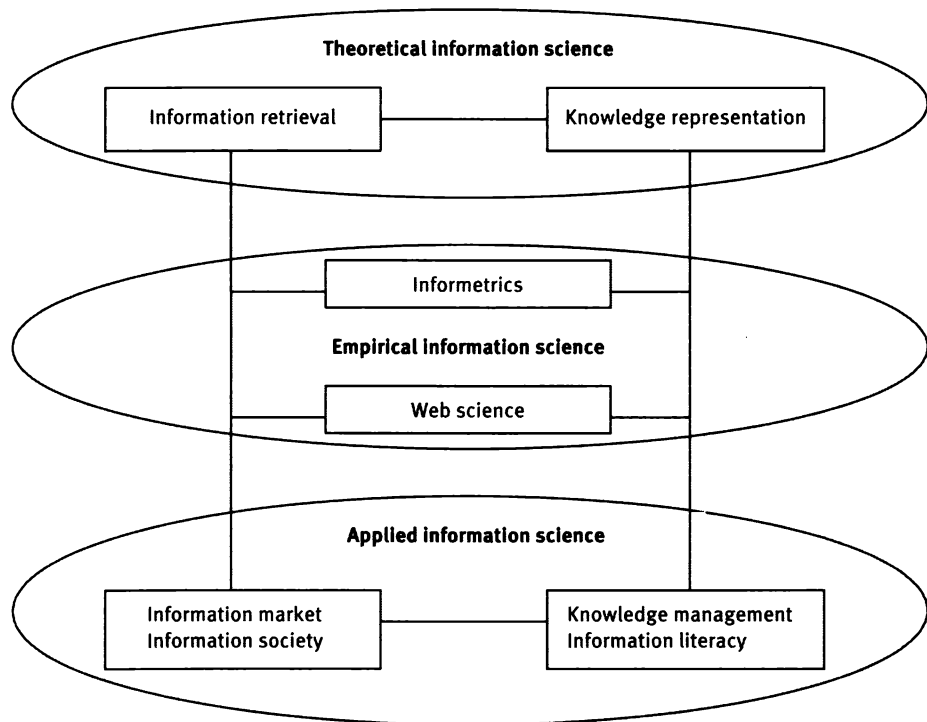


Figure A.1.1: Information Science and its Sub-Disciplines.

Information Science and Its Sub-Disciplines

Information science comprises a spectrum of five sub-disciplines:

- Information Retrieval,
- Knowledge Representation,
- Knowledge Management and Information Literacy,
- Research into the Information Society and Information Markets,
- Informetrics, including Web Science.

A central role is occupied by *Information Retrieval* (Baeza-Yates & Ribeiro-Neto, 2011; Manning, Raghavan, & Schütze, 2008; Stock, 2007), which is the science and engi-

neering of search engines (Croft, Metzler, & Strohman, 2010). Information retrieval investigates not only the technical systems, but also the information needs of the people (Cole, 2012) who use those systems (Ingwersen & Järvelin, 2005). The fundamental question of information retrieval is: how can one create technologically ideal retrieval systems in a user-optimized way?

Knowledge Representation addresses surrogates of knowledge and of documents in information systems. The theme is data describing documents and knowledge—so-called “metadata”—, such as a catalog card in a library (Taylor, 1999). Added to this are methods and tools for use during indexing and summarization (Chu, 2010; Lancaster, 2003; Stock & Stock, 2008). Essential tools are Knowledge Organization Systems (KOSs), which provide concepts and designations for indexing and retrieval. Here, the fundamental question is: how can knowledge in digital systems be represented, organized and condensed in such a way that it can be retrieved as easily as possible?

One subfield on the application side is *Knowledge Management* (Nonaka & Takeuchi, 1995), which focuses on the distribution and sharing of in-company knowledge as well as the integration of external knowledge into the corporate information systems. Another application of information science is research on *Information Literacy* (Eisenberg & Berkowitz, 1990), which means the use of (some) results of information science in everyday life, the workplace (Bruce, 1999), school (Chu, 2009) and university education (Johnston & Webber, 2003).

Research into the *Information Markets* (Linde & Stock, 2011) and the *Information Society* (Castells, 1996; Cronin, 2008; Webster, 1995) is particularly important with regard to the knowledge society. These research endeavors of information science have an extensive subject area, comprising everything from a country’s information infrastructure up to the network economics, by way of the industry of information service providers.

Information science proceeds empirically—where possible—and analyzes its objects via quantitative methods. *Informetric Research* (Weber & Stock, 2006; Wolfram, 2003) comprises webometrics (research into the World Wide Web; Thelwall, 2009), scientometrics (research into the information processes in science and medicine; van Raan, 1997; Haustein, 2012), “altmetrics” as an “alternative” metrics (informetrics with regard to services in Web 2.0; Priem & Hemminger, 2010) as well as patent informetrics (Schmitz, 2010). Empirical research is also performed during the evaluation of information systems (Voorhees, 2002) as well as during user and information needs analyses (Wilson, 2000). Additionally, informetrics attempts to detect the regularities or even laws of information processes (Egghe, 2005). One cross-section of informetrics, called *Web Science*, concentrates on research into the World Wide Web, its users and their information needs (Berners-Lee et al., 2006).

Information science has strong connections to a broad range of other scientific fields. It seems to be an interdisciplinary science (Buckland, 2012, 5). It combines, among others, methods and results from computer science (e.g., algorithms of

information retrieval and knowledge representation), the social sciences (e.g., user research), the humanities (e.g., linguistics and philosophy, analyzing concepts and words), economics (e.g., endeavors on the information markets), business administration (e.g., corporate knowledge management), education (e.g., applying information services in e-learning), and engineering (e.g., the construction of search engines for the World Wide Web) (Wilson, 1996). However, information science is by no means a “mixture” of other sciences, but a science on its own right. The central concern of information science—according to Buckland (2012, 6)—is “cultural engagement”, or, more precisely, “enabling people to become better informed (learning, become more knowledgeable)” (Buckland, 2012, 5). The application of information science in the normal course of life is to secure information literacy for all people. Information literate people are able to inform other people adequately (to represent, to store and to supply documents and knowledge) and they are always able to be well informed (to search for and to retrieve relevant documents and knowledge).

Information Science as Basic and Applied Research

We regard information science as both basic and applied research. In economic terms, it is an “importer” of ideas from other disciplines and an “exporter” of original results to other scientific branches (Cronin & Pearson, 1990). Information science is a basic discipline (and so an exporter of ideas) for some subfields of IT, of library science, of science policy and of the economic sciences. For instance, information science develops an innovative ranking algorithm for search engines, using user research and conceptual analyses. Computer science then takes up the suggestions and develops a workable system from them. On the other hand, information science also builds upon the results of computer science and other disciplines (e.g. linguistics) and imports their ideas. Endeavors toward the lemmatization of terms within texts, for instance, are doomed without an exact knowledge of linguistic results. The import/export-ratio of information science varies according to the partner discipline (Larivière, Sugimoto, & Cronin, 2012, 1011):

Although LIS’s (LIS: Library and Information Science, A/N) import dependency has been steadily decreasing since the mid-1990s—from 3.5 to about 1.3 in 2010 (inset)—it still has a negative balance of trade with most fields. The fields with which LIS has a positive balance of trade are from the natural, mostly medical sciences. Several of the fields with which LIS has a negative balance of trade are from the social sciences and the humanities.

Information science nearly always—even in questions of basic research—looks toward technological feasibility, incorporating user or usage as a matter of principle. It locates its object of research by investigating existing systems (such as search engines, library catalogs, Web 2.0 services and other information service providers’ systems) or creating experimental systems. Information science either analyzes and

evaluates such systems or does scientific groundwork for them. Designers of information services have to consider the understanding of the systems' users, their background knowledge, tradition and language. Thus information science is by no means only a technology, but also analyzes cognitive processes of the users and other stakeholders (Ingwersen & Järvelin, 2005).

Methods of knowledge representation are regarded independently of their historical provenance, converging into a single repertoire of methods: knowledge representation typically spans computer science aspects (ontologies; Gómez-Pérez, Fernández-López, & Corcho, 2004; Gruber, 1993; Staab & Studer, Eds., 2009), originally library-oriented or documentary endeavors (nomenclatures, classifications, thesauri; Lancaster, 2003) as well as ad hoc developments in the collaborative Web (folksonomies; Peters, 2009). Knowledge representation is an important building stone on the way to the “Semantic Web” (Berners-Lee, Hendler, & Lassila, 2001) or the “Social Semantic Web” (Weller, 2010), respectively.

Information Science as the Science of Information Content

The fixed point of information science is information itself, i.e. the structured information content which expresses knowledge. According to Buckland (1991), “information” has three aspects of meaning, all of which are objects of information science:

- information as a process (one informs / is informed),
- information as knowledge (information transports knowledge),
- information as a thing (information is fixed in “informative things”, i.e. in documents).

Documents are both texts (Belkin & Robertson, 1976) and non-textual objects (Buckland, 1997) such as images, music and videos, but they are also objects in science and technology (e.g., chemical substances and compounds), economic objects (such as companies or markets), objects in museums and galleries, real-time facts (e.g., weather data) as well as people (Stock, Peters, & Weller, 2010). Besides documents which are created by machines (e.g., flight-tracking systems), the majority of documents include “human recorded information” (Bawden & Robinson, 2012, 4). Documents are very important for information science (Davis & Shaw, Eds., 2011, 15):

Before information science was termed *information science*, it was called *documentation*, and documents were considered the basic objects of study for the field.

Of less interest to information science are technological information processing (which is also an object of IT) and the organization of information activities toward the sale of content (these are subject to the economic sciences as well). Thus Rauch (1988, 26) defines our discipline via the concept of knowledge:

For *information science* ... information is *knowledge*. More precisely: knowledge that is needed in order to deal with problematic situations. Knowledge is thus *possible* information, in a way. Information is knowledge that becomes effective and relevant for actions.

Kuhlen (2004, 5) even discusses the term “knowledge science”, which in light of the sub-disciplines of “knowledge representation” and “knowledge management” would not be completely beside the point. Kuhlen, too, adheres to “information science”, but emphasizes its close connection to knowledge (Kuhlen, 2004, 6):

Information does not exist in itself. Information refers to knowledge. Information is generally understood to be a surrogate, a representation or manifestation of knowledge.

The specific content plays a subordinate role in information science; its objects are the structure and function of information and information processing.

(For information science) it is of no importance whether the object of discussion is a new insect, for example, or an advanced method of metal working,

Mikhailov, Cernyi and Gilyarevsky (1979, 45) write. One of the “traditional” definitions of information science comes from Borko (1968, 3):

Information science is that discipline that investigates the properties and behavior of information, the forces governing the flow of information, and the means of processing information for optimum accessibility and usability. It is concerned with that body of knowledge relating to the origination, collection, organization, storage, retrieval, interpretation, transmission, transformation, and utilization of information. This includes the investigation of information representation in both natural and artificial systems, the use of codes for efficient message transmission, and the study of information processing devices and techniques such as computers and their programming systems. ... It has a pure science component, which inquires into the subject without to its application, and an applied science component, which develops services and products.

“That definition has been quite stable and unvarying over at least the last 30 years,” is Bates’s (1999, 1044) comment on this concept definition. Borko’s “pure science” can be further subdivided—as we have already seen in Figure A.1.1—into a theoretical information science, which works out the fundamentals of information retrieval as well as of knowledge representation, and into an empirical information science that systematically studies information systems and users (who deal with information systems). Applied information science addresses the use of information in practice, i.e. on the information markets, within a company or administration in knowledge management and in everyday life as information literacy. One application of information science in information practice aims toward keeping all members of a company, college, city, society or humanity in general as ideally informed as possible: everyone, in the right place and at the right time, must be able to receive the right amount of

relevant knowledge, rendered comprehensibly, clearly structured and condensed to the fundamental quantity.

Whereas information science does have some broad theoretical research areas, as a whole the discipline is rather oriented towards application. Even though certain phenomena may not yet be entirely resolved on a theoretical basis, an information scientist will still go ahead and create workable systems. Search engines on the Internet provide an illustrative example. The fundamental elements of processing queries and documents are nowhere exhaustively discussed and resolved in the theory of information retrieval; even the concept of relevance, which ought to be decisive for the relevance ranking being used, is anything but clear. Yet, for all that, the search engines built on this basis operate very successfully. Looking back on the advent of information science, Henrichs (1997, 945) describes the initial sparks for the discipline, which almost exclusively stem from information practice:

The practice of modern specialist information ... undoubtedly has a significant chronological advantage over its theory. Hence, in the beginning—and there can be no doubting this fact in this country (Germany, A/N), at least—there was practice.

As early as 1931, Ranganathan formulated five “laws” of library science. For Ranganathan, these are rules describing the processing of books in libraries. For today’s information science, the information content—whether fixed in books, websites or anywhere else—is the crucial aspect:

1st Law: Information is for use.

2nd Law: Every user his or her information.

3rd Law: Every information its user.

4th Law: Save the time of the user!

5th Law: Information practice and information science are growing organisms.

The first four rules refer to the aspect of practice, which always comes to bear on information science. The fifth “Law” states that the development of information practice and information science is not over yet by a long shot, and is constantly evolving.

A Short History of Information Science and Its Sub-Disciplines

The history of searching and finding knowledge begins during the period in which human beings first began to systematically solve problems. “As we strive to better understand the world that surrounds us, and to control it, we have a voracious appetite for information,” Norton (2000, 4) points out. The history of information science as a scientific discipline, however, goes back to the 1950s at its earliest (Bawden & Robinson, 2012, 8). The term “information science” was coined by Jason Farradane in the 1950s (Shapiro, 1995). Information science finally succeeded in establishing itself,

particularly in the USA, the United Kingdom and the former socialist countries, at the end of the 1960s (Schrader, 1984).

Information science has deep historical roots accented with significant controversy and conflicting views. The concepts of this science may be at the heart of many disciplines, but the emergence of a specific discipline of information science has been limited to the twentieth century (Norton 2000, 3).

In its “prehistory” and history, we can separate five strands that each deal with partial aspects of information science.

Information Retrieval

The retrieval of information has been discussed ever since there have been libraries. Systematically organized libraries help their users to purposefully retrieve the information they desire. Let the reader imagine being in front of a well-arranged bookshelf, and discovering a perfectly relevant work at a certain place (either via the catalog or by browsing). Apart from this direct hit, there will be further relevant books to its left and right, the significance of which to the topic of interest will diminish the further the user moves away from the center. The terms “catalog” (or, today, “search engine”), “bookshelf” (or “hit list”), “browsing”, “good arrangement” and “relevance” give us the basic concepts of information retrieval. The basic technology of information retrieval did not change much until the days of World War II. The advent of computers, however, changed the situation dramatically. Entirely new ways of information retrieval opened up. Experiments were made using computers as knowledge stores and elaborate retrieval languages were developed (Moovers, 1952; Luhn, 1961; Salton & McGill, 1983). Experimental retrieval systems were created in the 1960s, while commercial systems with specialist information have existed since the 1970s. Retrieval research reached its apex with the Internet search engines of the 1990s.

Knowledge Representation

What is a “good arrangement” of information? How can I represent knowledge ideally, i.e. condense it and make it retrievable via knowledge organization systems? This strand of development, too, stems from the world of libraries; it is closely related to information retrieval. Here, organization systems are at the forefront of the debate. Early evidence of such a system includes the systematics of the old library of Alexandria. Research into knowledge representation was boosted by the Decimal Classification developed by Mevil Dewey in the last quarter of the 19th century (Dewey, 1876), which was then developed further by Paul Otlet and Henri La Fontaine, finally leading to the plan of a collection and representation of world knowledge (Otlet, 1934). Over the course of the 20th century, various universal and specialized classification systems

were developed. With the triumph of computers in information practice, information scientists created new methods of knowledge representation (such as the thesaurus) as well as technologies for automatically indexing documents.

Knowledge Management and Information Literacy

In economics and business administration, information has long been discussed in the context of entrepreneurial decisions. Information always proves imperfect, and so becomes a motor for innovative competition. With conceptions for learning organizations, and, later, knowledge management (Nonaka & Takeuchi, 1995), the subject area of industrial economics and of business administration on the topic of information has broadened significantly from around 1980 onward. The objective is to share and safeguard in-company knowledge (Probst, Raub, & Romhardt, 2000) and to integrate external knowledge into an organization (Stock, 2000).

Information literacy has two main threads, namely retrieval skills and skills of creation and representation of knowledge. Retrieval literacy has its roots in library instruction and was introduced by librarians in the 1970s and 1980s (Eisenberg & Berkowitz, 1990). Since 1989, there exists a standard definition of information literacy, which was formulated by the American Library Association (Presidential Committee on Information Literacy, 1989). With the advent of Web 2.0 services, a second information literacy thread came into life: the skills of creating documents (e.g., videos for publishing on YouTube) and of representing them thematically (e.g., with tags).

Information Markets and Information Society

With the “knowledge industry” (Machlup, 1962), the “information economy” (Porat, 1977), the “postindustrial society” (Bell, 1973), the “information society” (Webster, 1995) or the “network society” (Castells, 1996), the focus on knowledge has for around 50 years been shaped by a sociological and economic perspective. In economics, information asymmetries (Akerlof, 1970) are discovered and measures to counteract the unsatisfactory aspects of markets with information asymmetries are developed. Thus, the sellers of digital information know a lot more about the quality of their product than their customers do, for example.

Originally, the information market was conceived as a very broad approach that led to the construction of a fourth economic factor (besides labor, capital and land) as well as to a fourth economic sector (besides agriculture, industry and services). Subsequently, the research area of the information markets was reduced to the sphere of digital information, which is transmitted via networks—the Internet in particular, but also mobile telephone networks (Linde & Stock, 2011). Research into the information, knowledge or network society thus discusses a very large area. Its subjects reach from “informational cities” as prototypes of cities in the knowledge society (Stock, 2011)

up to the “digital divide”, which is the result of social inequalities in the knowledge society (van Dijk & Hacker, 2003).

Informetrics and Web Science

Starting in the second quarter of the 20th century, researchers discovered that the distribution of information follows certain regularities. Studies of ranking distributions—e.g. of authors in a discipline according to their numbers of publications—led to laws, of which those formulated by Samuel C. Bradford, Alfred J. Lotka or George K. Zipf have become classics (Egghe, 2005). Chronological distributions of information are described via the concept of “half-life”. Since the 1970s, informetrics has established itself as information science’s measurement method. Some much-noticed fields of empirical information science include scientometrics and patentometrics, since informetrics allows us to measure scientific and technological information. This leads to the possibility of making statements about the “quality” of research results, and furthermore about the performance and influence of their developers and discoverers.

From around 2005 onward, an interdisciplinary research field into the World Wide Web, called Web Science, has emerged (Hendler et al., 2008). Information science takes part in this endeavor via webometrics, altmetrics, user research as well as information needs research.

Information Science and Its Neighbors

In his classification of the position of information science, Saracevic (1999, 1052) emphasizes its relations to other disciplines. Among these, he assigns particular importance to the role of information technology and the information society in interdisciplinary work.

First, information science is interdisciplinary in nature; however, the relations with various disciplines are changing. The interdisciplinary evolution is far from over.

Second, information science is inexorably connected to information technology. A technological imperative is compelling and constraining the evolution of information science (...).

Third, information science is, with many other fields, an active participant in the evolution of the information society. Information science has a strong social and human dimension, above and beyond technology.

Information science actively communicates with other scientific disciplines. It is closely related to the following:

- Computer Science,
- Economics,
- Library Science,

- (Computational) Linguistics,
- Pedagogy,
- Science of Science.

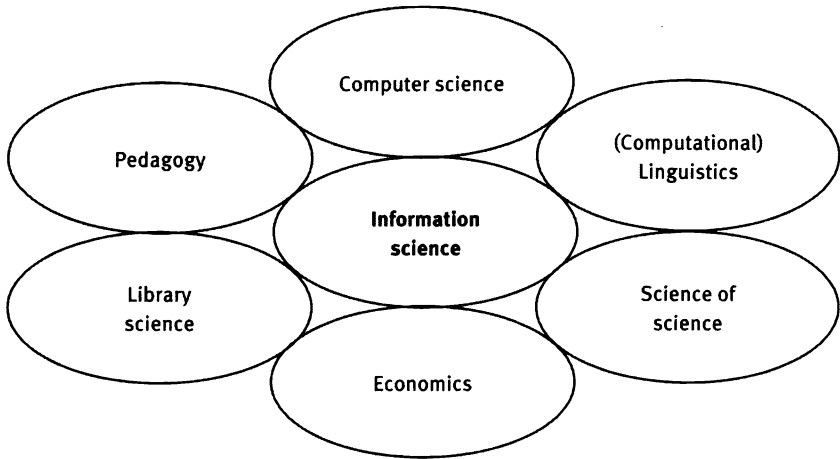


Figure A.1.2: Information Science and its Neighboring Disciplines.

Computer science provides the technological basis for information science applications. Without it, there would be no Internet and no commercial information industry. Some common research objects include information retrieval, aspects of knowledge representation (e.g., ontologies) and aspects of empirical information science, such as the evaluation of information systems. Computer science is more interested in the technology of information processing, information science in the processing of content. Unifying both perspectives suggests itself almost immediately. It has become common practice to refer to the combination between the two as Computer and Information Science—or “CIS”.

On the level of enterprises, the streams of information and communication between employees as well as with their environment are of extreme importance. Organizing both information technology and the usage thereof are tasks of information systems research and of information management. Organizing the in-company knowledge base as well as the sharing and communication of information fall within the domain of knowledge management. Information management, information systems research and knowledge management together form the corporate information economy. On the industry level, we regard information as a product. Electronic information services, search engines, Web portals etc. serve to locate a market for the product called “knowledge”. But this product has its idiosyncrasies, e.g. it can be copied at will, and many consumers regard it as a free common good. On the business level, lastly, network economics discusses the specifics of networks—of which the

Internet is one—and information economics asks about the buyer’s and the seller’s respective standards of knowledge, which already lays bare some considerable asymmetries.

The object of library science is the empirical and theoretical analysis of specific activities; among these are the collection, conservation, provision and evaluation of documents and the knowledge fixed therein. Its tools are elaborate systems for the formal and content-oriented processing of information. Topics like the creation of classification systems or information dissemination were common property of this discipline even before the term “information science” existed. This close link facilitates—especially in the United States—the development of approaches toward treating information science and library science as a single aggregate discipline, called “LIS” (Library and Information Science) (Milojević, Sugimoto, Yan, & Ding, 2011).

Search and retrieval are mainly performed via language; the knowledge—aside from some exceptions (such as images or videos)—is fixed linguistically. Computational linguistics and general linguistics are both so important for information science that relevant aspects from the different areas have been merged into a separate discipline, called information linguistics.

Pedagogy sometimes works with tools and services of Web 2.0, e.g. wikis or blogs. Additionally, in educational science learning management systems and e-portfolio systems are of great interest. Mainly in the aspects of e-learning and blended learning (Beutelspacher & Stock, 2011), there are overlaps with information science.

The analysis of scientific communication provided by empirical information science has made it possible to describe individual scientists, institutes, journals, even cities and countries. Such material is in strong demand in science of science and science policy. These results help sociology of science in researching for communities of scientists. History of science gets empirical historical source material. Science of science can check its hypotheses. And finally, science evaluation and science policy are provided with decision-relevant pointers toward the evaluation of scientific institutions.

Conclusion

-
- The object of information science is the representation, storage and supply as well as the search for and retrieval of relevant (predominantly digital) documents and knowledge.
 - Information science consists of five main sub-disciplines: (1) information retrieval (science and engineering of search engines), (2) knowledge representation (science and engineering of the storage and representation of knowledge), (3) informetrics (including all metrics applied in information science), (4) endeavors on the information markets and the information society (researches in economics and sociology as parts of applied information science), and (5) efforts on knowledge management as well as on information literacy (researches in business administration and education).

- Some application cases of information science research results are search engines, Web 2.0 services, library catalogs, digital libraries, specialist information service suppliers as well as information systems in corporate knowledge management.
 - The application of information science in the normal course of life is to secure information literacy for all people.
 - Historically speaking, five main lines of development can be detected in information science: information retrieval, knowledge representation, informetrics, information markets / information society and knowledge management / information literacy. Information science has begun its existence as an independent discipline in the 1950s, and is deemed to have been firmly established since around 1970.
 - Information science works between disciplines and has significant intersections with computer science, with economics, with library science, with linguistics, with pedagogy and with science of science.
-

Bibliography

- Akerlof, G.A. (1970). The market for “lemons”. Quality, uncertainty, and the market mechanism. *Quarterly Journal of Economics*, 84(3), 488-500.
- Baeza-Yates, R., & Ribeiro-Neto, B. (2011). *Modern Information Retrieval. The Concepts and Technology behind Search*. 2nd Ed. Harlow: Addison-Wesley.
- Bates, M.J. (1999). The invisible substrate of information science. *Journal of the American Society for Information Science*, 50(12), 1043-1050.
- Bawden, D., & Robinson, L. (2012). *Introduction to Information Science*. London: Facet.
- Belkin, N.J. (1978). Information concepts for information science. *Journal of Documentation*, 34(1), 55-85.
- Belkin, N.J., & Robertson, S.E. (1976). Information science and the phenomenon of information. *Journal of the American Society for Information Science*, 27(4), 197-204.
- Bell, D. (1973). *The Coming of the Post-Industrial Society. A Venture in Social Forecasting*. New York, NY: Basic Books.
- Berners-Lee, T., Hall, W., Hendler, J.A., O'Hara, K., Shadbolt, N., & Weitzner, D.J. (2006). A framework for web science. *Foundations and Trends in Web Science*, 1(1), 1-130.
- Berners-Lee, T., Hendler, J.A., & Lassila, O. (2001). The semantic Web. *Scientific American*, 284(5), 28-37.
- Beutelspacher, L., & Stock, W.G. (2011). Construction and evaluation of a blended learning platform for higher education. In R. Kwan, C. McNaught, P. Tsang, F.L. Wang, & K.C. Li (Eds.), *Enhanced Learning through Technology. Education Unplugged. Mobile Technologies and Web 2.0* (pp. 109-122). Berlin, Heidelberg: Springer. (Communications in Computer and Information Science; 177).
- Borko, H. (1968). Information science: What is it? *American Documentation*, 19(1), 3-5.
- Bruce, C.S. (1999). Workplace experiences of information literacy. *International Journal of Information Management*, 19(1), 33-47.
- Buckland, M.K. (1991). Information as thing. *Journal of the American Society for Information Science*, 42(5), 351-360.
- Buckland, M.K. (1997). What is a “document”? *Journal of the American Society for Information Science*, 48(9), 804-809.
- Buckland, M.K. (2012). What kind of science *can* information science be? *Journal of the American Society for Information Science and Technology*, 63(1), 1-7.

- Castells, M. (1996). *The Rise of the Network Society*. Malden, MA: Blackwell.
- Catts, R., & Lau, J. (2008). *Towards Information Literacy Indicators*. Paris: UNESCO.
- Chu, H. (2010). *Information Representation and Retrieval in the Digital Age*. 2nd Ed. Medford, NJ: Information Today.
- Chu, S.K.W. (2009). Inquiry project-based learning with a partnership of three types of teachers and the school librarian. *Journal of the American Society for Information Science and Technology*, 60(8), 1671-1686.
- Cole, C. (2012). *Information Need. A Theory Connecting Information Search to Knowledge Formation*. Medford, NJ: Information Today.
- Croft, W.B., Metzler, D., & Strohman, T. (2010). *Search Engines. Information Retrieval in Practice*. Boston, MA: Addison Wesley.
- Cronin, B. (2008). The sociological turn in information science. *Journal of Information Science*, 34(4), 465-475.
- Cronin, B., & Pearson, S. (1990). The export of ideas from information science. *Journal of Information Science*, 16(6), 381-391.
- Davis, C.H., & Shaw, D. (Eds.) (2011). *Introduction to Information Science and Technology*. Medford, NJ: Information Today.
- Dewey, M. (1876). *A Classification and Subject Index for Cataloguing and Arranging the Books and Pamphlets of a Library*. Amherst, MA (anonymous).
- Egghe, L. (2005). *Power Laws in the Information Production Process*. Lotkaian Informetrics. Amsterdam: Elsevier.
- Eisenberg, M.B., & Berkowitz, R.E. (1990). *Information Problem-Solving. The Big Six Skills Approach to Library & Information Skills Instruction*. Norwood, NJ: Ablex.
- Gómez-Pérez, A., Fernández-López, M., & Corcho, O. (2004). *Ontological Engineering*. London: Springer.
- Gruber, T.R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2), 199-220.
- Haustein, S. (2012). *Multidimensional Journal Evaluation. Analyzing Scientific Periodicals beyond the Impact Factor*. Berlin, Boston, MA: De Gruyter Saur. (Knowledge & Information. Studies in Information Science.)
- Hendler, J.A., Shadbolt, N., Hall, W., Berners-Lee, T., & Weitzner, D. (2008). Web science. An interdisciplinary approach to understanding the web. *Communications of the ACM*, 51(7), 60-69.
- Henrichs, N. (1997). *Informationswissenschaft*. In W. Rehfeld, T. Seeger, & D. Strauch (Eds.), *Grundlagen der praktischen Information und Dokumentation* (pp. 945-957). 4th Ed. München: Saur.
- Ingwersen, P., & Järvelin, K. (2005). *The Turn. Integration of Information Seeking and Retrieval in Context*. Dordrecht: Springer.
- Johnston, B., & Webber, S. (2003). Information literacy in higher education. A review and case study. *Studies in Higher Education*, 28(3), 335-352.
- Kuhlen, R. (2004). *Information*. In R. Kuhlen, T. Seeger, & D. Strauch (Eds.), *Grundlagen der praktischen Information und Dokumentation* (pp. 3-20). 5th Ed. München: Saur.
- Lancaster, F.W. (2003). *Indexing and Abstracting in Theory and Practice*. 3rd Ed. Champaign, IL: University of Illinois.
- Larivière, V., Sugimoto, C.R., & Cronin, B. (2012). A bibliometric chronicling of library and information science's first hundred years. *Journal of the American Society for Information Science and Technology*, 63(5), 997-1016.
- Linde, F., & Stock, W.G. (2011). *Information Markets. A Strategic Guideline for the I-Commerce*. Berlin, New York, NY: De Gruyter Saur. (Knowledge & Information. Studies in Information Science.)

- Luhn, H.P. (1961). The automatic derivation of information retrieval encodements from machine-readable texts. In A. Kent (Ed.), *Information Retrieval and Machine Translation*, Vol. 3, Part 2 (pp. 1021-1028). New York, NY: Interscience.
- Machlup, F. (1962). *The Production and Distribution of Knowledge in the United States*. Princeton, NJ: Princeton University Press.
- Manning, C.D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge: Cambridge University Press.
- Mikhailov, A.I., Cernyi, A.I., & Gilyarevsky, R.S. (1979). *Informatik*. *Informatik*, 26(4), 42-45.
- Milojević, S., Sugimoto, C.R., Yan, E., & Ding, Y. (2011). The cognitive structure of library and information science. Analysis of article title words. *Journal of the American Society for Information Science and Technology*, 62(10), 1933-1953.
- Moors, C.N. (1952). Information retrieval viewed as temporal signalling. In *Proceedings of the International Congress of Mathematicians*. Cambridge, Mass., August 30 – September 6, 1950. Vol. 1 (pp. 572-573). Providence, RI: American Mathematical Society.
- Nonaka, I., & Takeuchi, H. (1995). *The Knowledge-Creating Company. How Japanese Companies Create the Dynamics of Innovation*. Oxford: Oxford University Press.
- Norton, M.J. (2000). *Introductory Concepts in Information Science*. Medford, NJ: Information Today.
- Otlet, P. (1934). *Traité de Documentation*. Bruxelles: Mundaneum.
- Peters, I. (2009). *Folksonomies. Indexing and Retrieval in Web 2.0*. Berlin: De Gruyter Saur. (Knowledge & Information. Studies in Information Science.)
- Porat, M.U. (1977). *Information Economy*. 9 Vols. (OT Special Publication 77-12[1] – 77-12[9]). Washington, DC: Office of Telecommunication.
- Presidential Committee on Information Literacy (1989). *Final Report*. Washington, DC: American Library Association / Association for College & Research Libraries.
- Priem, J., & Hemminger, B. (2010). *Scientometrics 2.0. Toward new metrics of scholarly impact on the social Web*. *First Monday*, 15(7).
- Probst, G.J.B., Raub, S., & Romhardt, K. (2000). *Managing Knowledge. Building Blocks for Success*. Chichester: Wiley.
- Ranganathan, S.R. (1931). *The Five Laws of Library Science*. Madras: Madras Library Association, London: Edward Goldston.
- Rauch, W. (1988). *Was ist Informationswissenschaft?* Graz: Kienreich.
- Salton, G., & McGill, M.J. (1983). *Introduction to Modern Information Retrieval*. New York, NY: McGraw-Hill.
- Saracevic, T. (1999). Information Science. *Journal of the American Society for Information Science*, 50(12), 1051-1063.
- Schmitz, J. (2010). *Patentinformatik. Analyse und Verdichtung von technischen Schutzrechtsinformationen*. Frankfurt/Main: DGI.
- Schrader, A.M. (1984). In search of a name. *Information science and its conceptual antecedents*. *Library and Information Science Research*, 6(3), 227-271.
- Shapiro, F.R. (1995). Coinage of the term *information science*. *Journal of the American Society for Information Science*, 46(5), 384-385.
- Staab, S., & Studer, R. (Eds.) (2009). *Handbook on Ontologies*. Dordrecht: Springer.
- Stock, W.G. (2000). *Informationswirtschaft. Management externen Wissens*. München: Oldenbourg.
- Stock, W.G. (2007). *Information Retrieval. Informationen suchen und finden*. München: Oldenbourg.
- Stock, W.G. (2011). Informational cities. Analysis and construction of cities in the knowledge society. *Journal of the American Society for Information Science and Technology*, 62(5), 963-986.
- Stock, W.G., Peters, I., & Weller, K. (2010). Social semantic corporate digital libraries. Joining knowledge representation and knowledge management. *Advances in Librarianship*, 32, 137-158.

- Stock, W.G., & Stock, M. (2008). *Wissensrepräsentation. Informationen auswerten und bereitstellen*. München: Oldenbourg.
- Taylor, A.G. (1999). *The Organization of Information*. Englewood, CO: Libraries Unlimited.
- Thelwall, M. (2009). *Introduction to Webometrics. Quantitative Web Research for the Social Sciences*. San Rafael, CA: Morgan & Claypool.
- van Dijk, J., & Hacker, K. (2003). The digital divide as a complex and dynamic phenomenon. *The Information Society*, 19(4), 315-326.
- van Raan, A.F.J. (1997). Scientometrics. State-of-the-art. *Scientometrics*, 38(1), 205-218.
- Voorhees, E.M. (2002). The philosophy of information retrieval evaluation. *Lecture Notes in Computer Science*, 2406, 355-370.
- Weber, S., & Stock, W.G. (2006). Facets of informetrics. *Information – Wissenschaft und Praxis*, 57(8), 385-389.
- Webster, F. (1995). *Theories of the Information Society*. London: Routledge.
- Weller, K. (2010). *Knowledge Representation in the Social Semantic Web*. Berlin: De Gruyter Saur. (Knowledge & Information. Studies in Information Science.)
- Wilson, P. (1996). The future of research in our field. In J.L. Olaisen, E. Munch-Petersen, & P. Wilson (Eds.), *Information Science. From the Development of the Discipline to Social Interaction* (pp. 319-323). Oslo: Scandinavian University Press.
- Wilson, T.D. (2000). Human information behavior. *Informing Science*, 3(2), 49-55.
- Wolfram, D. (2003). *Applied Informetrics for Information Retrieval Research*. Westport, CT: Libraries Unlimited.
- Yan, X.S. (2011). Information science. Its past, present and future. *Information*, 2(3), 510-527.